Comment on 'Estimation and prediction of the HIV-AIDS-epidemic under conditions of HAART using mixtures of incubation time distributions' (letter)

Heisterkamp, S.H.; de Vries, R.; Sprenger, H.G.; Hubben, G.A.A.; Postma, M.J.

Published in:
Statistics in Medicine

Publication date:
2008

STATISTICS IN MEDICINE
Statist. Med. 2008; **27**:2035–2036
Published online 12 September 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/sim.3050

LETTER TO THE EDITOR

Comment on 'Estimation and prediction of the HIV-AIDS-epidemic under conditions of HAART using mixtures of incubation time distributions'

by S. H. Heisterkamp, R. de Vries, H. G. Sprenger, G. A. A. Hubben, M. J. Postma, *Statistics in Medicine*, DOI: 10.1002/sim.2974

From: *Saralees Nadarajah*
School of Mathematics,
University of Manchester,
Manchester M60 1QD, *U.K.*

The recent paper by Heisterkamp *et al.* [1] derived the incubation time distributions for HIV treated with the HAART regimen from a cohort study. This entailed the study of the convolution $Z = \sum_{i=1}^{k} X_i$, where X_i, $i = 1, 2, \ldots, k$ are independent Gumbel random variables specified by the cumulative distribution functions (cdfs):

$$F_{X_i}(x) = \exp\left\{-\exp\left(-\frac{x-\alpha_i}{\beta}\right)\right\}$$

for $-\infty < x < \infty$, $-\infty < \alpha_i < \infty$ and $\beta > 0$. Heisterkamp *et al.* stated that the distribution of Z looks 'intractable for convolutions $k > 2$', and proposed an approximation based on the fact 'that for k not too large the form of the distribution will remain Gumbel'. This fact does not appear to be true even for the case $k = 2$, see Figure 1. Heisterkamp *et al.* also derived an elementary expression for the cdf of Z for the very restrictive case $k = 2$ and $\alpha_1 = \alpha_2$, a result that is not at all new: it follows from the well-known fact that the difference between two independent and identically distributed Gumbel random variables is a logistic random variable, see Johnson *et al.* [2].

The main point of this letter is to point out that explicit expressions for the probability density function (pdf) of $Z = \sum_{i=1}^{k} X_i$ have been derived in the most general form when X_i are specified by the cdfs

$$F_{X_i}(x) = \exp\left\{-\exp\left(-\frac{x-\alpha_i}{\beta_i}\right)\right\}$$

for $-\infty < x < \infty$, $-\infty < \alpha_i < \infty$ and $\beta_i > 0$. For the case $k = 2$, Nadarajah [3, Theorem 1] derived an expression for the pdf of Z in terms of hypergeometric functions. For $k > 2$, Nadarajah [4, Theorem 1 and Corollary 1] derived an expression for the pdf in terms of the Meijer G function. The hypergeometric and Meijer G functions are well known and well established (see Sections 9.23

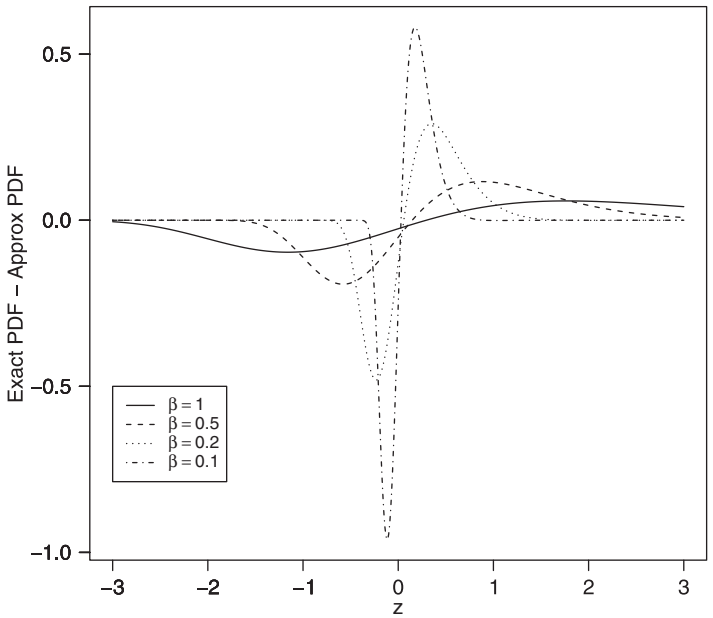

Figure 1. Difference between the exact and approximate pdfs of $X_1 + X_2$ for $\alpha_1 = 0$, $\alpha_2 = 0$ and $\beta = 0.1, 0.2, 0.5, 1$. The approximation proposed by Heisterkamp *et al.* is used.

and 9.3 of Gradshteyn and Ryzhik [5]) and in-built numerical routines for them are widely available (see Maple and Mathematica).

REFERENCES

1. Heisterkamp SH, de Vries R, Sprenger HG, Hubben GAA, Postma MJ. Estimation and prediction of the HIV-AIDS-epidemic under conditions of HAART using mixtures of incubation time distributions. *Statistics in Medicine* 2007; DOI: 10.1002/sim.2974.
2. Johnson NL, Kotz S, Balakrishnan N. *Continuous Univariate Distributions*, vol. 2 (2nd edn). Wiley: New York, 1995.
3. Nadarajah S. Linear combination of Gumbel random variables. *Stochastic Environmental Research and Risk Assessment* 2007; **21**:283–286.
4. Nadarajah S. Exact distribution of the linear combination of p Gumbel random variables. *International Journal of Computer Mathematics* 2007, in press.
5. Gradshteyn IS, Ryzhik IM. *Table of Integrals, Series, and Products* (6th edn). Academic Press: San Diego, CA, 2000.

